

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/005,3186
Source: EFW16
Date Processed by STIC: 8-2-05

ENTERED

**CRF Errors Edited by the STIC Systems
Branch**

Serial Number: 09/005,318G

CRF Edit Date: 8/2/05
Edited by: KQI

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

✓ ___ Corrected the SEQ ID NO. Sequence numbers edited were:

13

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

___ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

___ Moved responses to same line as heading/numeric identifier, specifically:

✓ ___ Other:

Corrected amino Acid
numbering for seq ID # 113.



IFW16

RAW SEQUENCE LISTING

DATE: 08/02/2005

PATENT APPLICATION: US/09/005,318G

TIME: 11:19:27

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

3 <110> APPLICANT: HEIN, MICH B.
 4 HIATT, ANDREW C.
 5 FITCHEN, JOHN H.
 7 <120> TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
 9 <130> FILE REFERENCE: 040989/283662
 11 <140> CURRENT APPLICATION NUMBER: 09/005,318G
 12 <141> CURRENT FILING DATE: 1998-01-09
 14 <150> PRIOR APPLICATION NUMBER: 08/782,481
 15 <151> PRIOR FILING DATE: 1997-01-10
 17 <150> PRIOR APPLICATION NUMBER: 09/005,167
 18 <151> PRIOR FILING DATE: 1998-01-09
 20 <160> NUMBER OF SEQ ID NOS: 140
 22 <170> SOFTWARE: PatentIn Ver. 2.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 137
 26 <212> TYPE: PRT
 27 <213> ORGANISM: Homo sapiens
 29 <400> SEQUENCE: 1
 30 Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys Ala
 31 1 5 10 15
 33 Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu Asp
 34 20 25 30
 36 Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg Glu
 37 35 40 45
 39 Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Pro Val Tyr His
 40 50 55 60
 42 Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu Asp
 43 65 70 75 80
 45 Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp Ser
 46 85 90 95
 48 Ala Thr Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Ala
 49 100 105 110
 51 Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr Ala
 52 115 120 125
 54 Leu Thr Pro Asp Ala Cys Tyr Pro Asp
 55 130 135
 58 <210> SEQ ID NO: 2
 59 <211> LENGTH: 135
 60 <212> TYPE: PRT
 61 <213> ORGANISM: Mus sp.
 63 <400> SEQUENCE: 2
 64 Gln Asp Glu Asn Glu Arg Ile Val Val Asp Asn Lys Cys Lys Cys Ala
 65 1 5 10 15

(pg. 6)

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/005,318G

DATE: 08/02/2005

TIME: 11:19:27

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

```

67 Arg Ile Thr Ser Arg Ile Ile Pro Ser Ala Glu Asp Pro Ser Gln Asp
68           20           25           30
70 Ile Val Glu Arg Asn Val Arg Ile Ile Val Pro Leu Asn Ser Arg Glu
71           35           40           45
73 Asn Ile Ser Asp Pro Thr Ser Pro Met Arg Thr Lys Pro Val Tyr His
74           50           55           60
76 Leu Ser Asp Leu Cys Lys Lys Cys Asp Thr Thr Glu Val Glu Leu Glu
77 65           70           75           80
79 Asp Gln Val Val Thr Ala Ser Gln Ser Asn Ile Cys Asp Ser Asp Ala
80           85           90           95
82 Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Asn Arg Val
83           100          105          110
85 Lys Leu Ser Tyr Arg Gly Gln Thr Lys Met Val Glu Thr Ala Leu Thr
86           115          120          125
88 Pro Asp Ser Cys Tyr Pro Asp
89           130          135
92 <210> SEQ ID NO: 3
93 <211> LENGTH: 137
94 <212> TYPE: PRT
95 <213> ORGANISM: Oryctolagus cuniculus
97 <400> SEQUENCE: 3
98 Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn Lys Cys Met Cys Thr Arg
99 1           5           10          15
101 Val Thr Ser Arg Ile Ile Pro Ser Thr Glu Asp Pro Asn Glu Asp Ile
102           20           25           30
104 Val Glu Arg Asn Ile Arg Ile Val Val Pro Leu Asn Asn Arg Glu Asn
105           35           40           45
107 Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg Asn Pro Val Tyr His Leu
108           50           55           60
110 Ser Asp Val Cys Lys Lys Cys Asp Pro Val Glu Val Glu Leu Glu Asp
111 65           70           75           80
113 Gln Val Val Thr Ala Thr Gln Ser Asn Ile Cys Asn Glu Asp Asp Gly
114           85           90           95
116 Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg Asn Lys Cys Tyr Thr Thr
117           100          105          110
119 Met Val Pro Leu Arg Tyr His Gly Glu Thr Lys Met Val Gln Ala Ala
120           115          120          125
122 Leu Thr Pro Asp Ser Cys Tyr Pro Asp
123           130          135
126 <210> SEQ ID NO: 4
127 <211> LENGTH: 136
128 <212> TYPE: PRT
129 <213> ORGANISM: Bos sp.
131 <400> SEQUENCE: 4
132 Glu Asp Glu Ser Thr Val Leu Val Asp Asn Lys Cys Gln Cys Val Arg
133 1           5           10          15
135 Ile Thr Ser Arg Ile Ile Arg Asp Pro Asp Asn Pro Ser Glu Asp Ile
136           20           25           30
138 Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Thr Arg Glu Asn

```

RAW SEQUENCE LISTING

DATE: 08/02/2005

PATENT APPLICATION: US/09/005,318G

TIME: 11:19:27

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

```

139          35          40          45
141 Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Glu Pro Lys Tyr Asn Leu
142          50          55          60
144 Ala Asn Leu Cys Lys Lys Cys Asp Pro Thr Glu Ile Glu Leu Asp Asn
145 65          70          75          80
147 Gln Val Phe Thr Ala Ser Gln Ser Asn Ile Cys Pro Asp Asp Asp Tyr
148          85          90          95
150 Ser Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Thr Leu
151          100          105          110
154 Val Pro Ile Thr His Arg Gly Val Thr Arg Met Val Lys Ala Thr Leu
155          115          120          125
157 Thr Pro Asp Ser Cys Tyr Pro Asp
158          130          135
161 <210> SEQ ID NO: 5
162 <211> LENGTH: 119
163 <212> TYPE: PRT
164 <213> ORGANISM: Rana sp.
166 <220> FEATURE:
167 <221> NAME/KEY: MOD_RES
168 <222> LOCATION: (47)
169 <223> OTHER INFORMATION: Variable amino acid
171 <220> FEATURE:
172 <221> NAME/KEY: MOD_RES
173 <222> LOCATION: (88)..(89)
174 <223> OTHER INFORMATION: Variable amino acid
176 <220> FEATURE:
177 <221> NAME/KEY: MOD_RES
178 <222> LOCATION: (91)
179 <223> OTHER INFORMATION: Variable amino acid
181 <400> SEQUENCE: 5
182 Glu Gln Glu Tyr Ile Leu Ala Asn Asn Lys Cys Lys Cys Val Lys Ile
183 1          5          10          15
185 Ser Ser Arg Phe Val Pro Ser Thr Glu Arg Pro Gly Glu Glu Ile Leu
186          20          25          30
W--> 188 Glu Arg Asn Ile Gln Ile Thr Ile Pro Thr Ser Ser Arg Met Xaa Ile
189          35          40          45
191 Ser Asp Pro Tyr Ser Pro Leu Arg Thr Gln Pro Val Tyr Asn Leu Trp
192          50          55          60
195 Asp Ile Cys Gln Lys Cys Asp Pro Val Gln Leu Glu Ile Gly Gly Ile
196 65          70          75          80
198 Pro Val Leu Ala Ser Gln Pro Xaa Xaa Ser Xaa Pro Asp Asp Glu Cys
199          85          90          95
201 Tyr Thr Thr Glu Val Asn Phe Lys Lys Lys Val Pro Leu Thr Pro Asp
202          100          105          110
204 Ser Cys Tyr Glu Tyr Ser Glu
205          115
208 <210> SEQ ID NO: 6
209 <211> LENGTH: 128
210 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 08/02/2005

PATENT APPLICATION: US/09/005,318G

TIME: 11:19:27

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

```

211 <213> ORGANISM: Lumbricus sp.
213 <400> SEQUENCE: 6
214 Asn Lys Cys Met Cys Thr Arg Val Thr Ala Arg Ile Arg Gly Thr Arg
215 1 5 10 15
217 Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Tyr Ile Arg Ile Asn Val
218 20 25 30
220 Pro Leu Lys Asn Arg Gly Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg
221 35 40 45
223 Asn Gln Pro Val Tyr His Leu Ser Pro Ser Cys Lys Lys Cys Asp Pro
224 50 55 60
226 Tyr Glu Asp Gly Val Val Thr Ala Thr Glu Thr Asn Ile Cys Tyr Pro
227 65 70 75 80
229 Asp Gln Gly Val Pro Gln Ser Cys Arg Asp Tyr Cys Pro Glu Leu Asp
230 85 90 95
232 Arg Asn Lys Cys Tyr Thr Val Leu Val Pro Pro Gly Tyr Thr Gly Glu
233 100 105 110
235 Thr Lys Met Val Gln Asn Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
236 115 120 125
239 <210> SEQ ID NO: 7
240 <211> LENGTH: 421
241 <212> TYPE: DNA
242 <213> ORGANISM: Homo sapiens
244 <220> FEATURE:
245 <221> NAME/KEY: CDS
246 <222> LOCATION: (1)..(414)
248 <220> FEATURE:
249 <221> NAME/KEY: sig_peptide
250 <222> LOCATION: (1)..(6)
252 <220> FEATURE:
253 <221> NAME/KEY: mat_peptide
254 <222> LOCATION: (7)..(414)
256 <400> SEQUENCE: 7
257 gat cag gaa gat gaa cgt att gtt ctg gtt gac aac aag tgc aag tgt 48
258 Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys
259 -1 1 5 10
261 gct cgt att act tct aga atc atc cgt agc tca gag gac cca aat gaa 96
262 Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu
263 15 20 25 30
265 gat ata gtc gaa cgt aac atc cgt atc atc gtc cca ctg aat aac cgg 144
266 Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg
267 35 40 45
269 gag aat atc tca gat cct aca agt ccg ttg cgc aca cgc ttc gta tac 192
270 Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr
271 50 55 60
273 cac ctg tca gat ctg tgt aag aag tgt gat cca aca gag gta gag ctg 240
274 His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu
275 65 70 75
277 gac aat cag ata gtc act gcg act caa agc aac att tgc gat gag gac 288
278 Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp

```

RAW SEQUENCE LISTING

DATE: 08/02/2005

PATENT APPLICATION: US/09/005,318G

TIME: 11:19:27

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

```

279      80      85      90
281 agc gct aca gaa acc tgc agc acc tac gat agg aac aaa tgc tac acg 336
282 Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr
283 95      100      105      110
285 gcc gtg gtt ccg ctc gtg tat ggt gga gag aca aaa atg gtg gaa act 384
286 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
287      115      120      125
289 gcc ctt acg ccc gat gca tgc tat ccg gac tgaattc 421
290 Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
291      130      135
295 <210> SEQ ID NO: 8
296 <211> LENGTH: 215
297 <212> TYPE: DNA
298 <213> ORGANISM: Homo sapiens
300 <220> FEATURE:
301 <221> NAME/KEY: CDS
302 <222> LOCATION: (1)..(213)
304 <400> SEQUENCE: 8
305 gat cag aag tgc aag tgt gct cgt att act tct aga atc atc cgt agc 48
306 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
307 1      5      10      15
309 tca gag gac cca aat gaa gat ata gtc gaa cgt aac atc cgt atc atc 96
310 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
311      20      25      30
313 gtc cca ctg aat aac cgg gag aat atc tca gat cct aca agt ccg ttg 144
314 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
315      35      40      45
317 cgc aca cgc ttc gta tac cac ctg tca gat ctg tgt aag aag gat gag 192
318 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Asp Glu
319      50      55      60
321 gac agc gct aca gaa acc tgc tg 215
322 Asp Ser Ala Thr Glu Thr Cys
323 65      70
327 <210> SEQ ID NO: 9
328 <211> LENGTH: 140
329 <212> TYPE: DNA
330 <213> ORGANISM: Homo sapiens
332 <400> SEQUENCE: 9
333 ctagaatcat ccgtagctca gaggacccaa atgaagatat agtcgaacgt aacatccgta 60
334 tcatcgtccc actgaataac cgggagaata tctcagatcc tacaagtccg ttgcgcacac 120
335 gcttcgtata ccacctgtca 140
338 <210> SEQ ID NO: 10
339 <211> LENGTH: 31
340 <212> TYPE: DNA
342 <213> ORGANISM: Homo sapiens
344 <400> SEQUENCE: 10
345 gatcagaagt gcaagtgtgc tcgtattact t 31
348 <210> SEQ ID NO: 11
349 <211> LENGTH: 44

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/005,318G

DATE: 08/02/2005
TIME: 11:19:28

Input Set : A:\PTO.KD.txt
Output Set: N:\CRF4\08022005\I005318G.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; Xaa Pos. 47,88,89,91
Seq#:118; Xaa Pos. 37,60,62
Seq#:123; Xaa Pos. 37,78,79,81

VERIFICATION SUMMARY

DATE: 08/02/2005

PATENT APPLICATION: US/09/005,318G

TIME: 11:19:28

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\08022005\I005318G.raw

L:188 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:32

M:341 Repeated in SeqNo=5

L:1864 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118 after pos.:32

M:341 Repeated in SeqNo=118

L:1965 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123 after pos.:32

M:341 Repeated in SeqNo=123

Raw Sequence Listing before editing,
for reference only



IFW16

RAW SEQUENCE LISTING

DATE: 07/29/2005

PATENT APPLICATION: US/09/005,318G

TIME: 13:31:40

Input Set : A:\283662 FINAL SEQLIST.txt

Output Set: N:\CRF4\07292005\I005318G.raw

3 <110> APPLICANT: HEIN, MICH B.
 4 HIATT, ANDREW C.
 5 FITCHEN, JOHN H.
 7 <120> TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
 9 <130> FILE REFERENCE: 040989/283662
 11 <140> CURRENT APPLICATION NUMBER: 09/005,318G
 12 <141> CURRENT FILING DATE: 1998-01-09
 14 <150> PRIOR APPLICATION NUMBER: 08/782,481
 15 <151> PRIOR FILING DATE: 1997-01-10
 17 <150> PRIOR APPLICATION NUMBER: 09/005,167
 18 <151> PRIOR FILING DATE: 1998-01-09
 20 <160> NUMBER OF SEQ ID NOS: 140
 22 <170> SOFTWARE: PatentIn Ver. 2.1

Does Not Comply
 Corrected Diskette Needed
 (Pg. 1-2)

ERRORED SEQUENCES

373 <210> SEQ ID NO: 13
 374 <211> LENGTH: 286
 375 <212> TYPE: DNA
 376 <213> ORGANISM: Homo sapiens
 378 <220> FEATURE:
 379 <221> NAME/KEY: CDS
 380 <222> LOCATION: (1)...(282)
 E--> 382 <400> SEQUENCE: 29-13

383	gac aac aag tgc aag tgt gct cgt att act tct aga atc atc cgt agc	48
384	Asp Asn Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser	
385	1 5 10 15	
387	tca gag gac cca aat gaa gat ata gtc gaa cgt aac atc cgt atc atc	96
388	Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile	
389	20 25 30	
391	gtc cca ctg aat aac cgg gag aat atc tca gat cct aca agt ccg ttg	144
392	Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu	
393	35 40 45	
395	cgc aca cgc ttc gta tac cac ctg tca gat ctg tgt aag aag tgt gat	192
396	Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Cys Asp	
397	50 55 60	
399	cca aca gag gta gag ctg gac aat cag ata gtc act gcg act caa agc	240
400	Pro Thr Glu Val Glu Leu Asp Asn Gln Ile Val Thr Ala Thr Gln Ser	
401	65 70 75 80	
403	aac att tgc gat gag gac agc gct aca gaa acc tgc tac tga	282
404	Asn Ile Cys Asp Glu Asp Ser Ala Thr Glu Thr Cys Tyr *	
405	85 90	

RAW SEQUENCE LISTING

DATE: 07/29/2005

PATENT APPLICATION: US/09/005,318G

TIME: 13:31:41

Input Set : A:\283662 FINAL SEQLIST.txt

Output Set: N:\CRF4\07292005\I005318G.raw

407 attc

286

1770 <210> SEQ ID NO: 113

1771 <211> LENGTH: 9

1772 <212> TYPE: PRT

1773 <213> ORGANISM: Artificial Sequence

1775 <220> FEATURE:

1776 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative

1777 peptide

1779 <400> SEQUENCE: 113

1780 Glu Gln Lys Leu Ile Ser Glu Asp Leu

E--> 1781 1 5

S

VERIFICATION SUMMARY

DATE: 07/29/2005

PATENT APPLICATION: US/09/005,318G

TIME: 13:31:42

Input Set : A:\283662 FINAL SEQLIST.txt

Output Set: N:\CRF4\07292005\I005318G.raw

L:188 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:32

M:341 Repeated in SeqNo=5

L:382 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:13 differs:29

L:1781 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:113

L:1864 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118 after pos.:32

M:341 Repeated in SeqNo=118

L:1965 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123 after pos.:32

M:341 Repeated in SeqNo=123